

Los indigenismos en la lexicografía guatemalteca decimonónica: el caso de los *Vicios del lenguaje y provincialismos de Guatemala* de Batres Jáuregui*

Indigenous American Words in Nineteenth-Century Guatemalan Lexicography: The Example of Batres Jáuregui's *Vicios del lenguaje y provincialismos de Guatemala*

José Luis Ramírez Luengo
Instituto Univ. Seminario Menéndez Pidal
Universidad Complutense de Madrid
joseluis.ramirezluengo@gmail.com
ORCID: 0000-0002-5564-2372

Resumen

Aunque el guatemalteco Antonio Batres Jáuregui constituye un nombre de primera importancia dentro de la etapa normativista de la lexicografía centroamericana, todavía no ha recibido, por parte de los investigadores, la atención que se les ha prestado a otros como, por ejemplo, Gagini o Membreño. En tal sentido, este trabajo pretende analizar los indigenismos recogidos en su *Vicios del lenguaje y provincialismos de Guatemala* (Ciudad de Guatemala: Encuadernación y Tipografía Nacional, 1892) con el propósito de conocer la incidencia que tienen estas voces en la obra mencionada; de este modo, se atenderá a cuestiones como el peso de tales vocablos en el total de los elementos censurados, su origen etimológico, los campos léxicos a los que pertenecen y los juicios de valor explícitos que a este respecto aparecen en las páginas del volumen mencionado, en un intento de comprender de manera más profunda no solo las ideas lingüísticas que sustentan la reflexión de Batres Jáuregui, sino también, de forma más general, la valoración que de las lenguas autóctonas hace la lexicografía centroamericana de finales del siglo XIX.

PALABRAS CLAVE: historia del español de América, lexicografía, siglo XIX, Guatemala, indigenismo.

* Esta publicación forma parte del proyecto de I+D+i PID2020-117659GB-I00, financiado por MCIN/ AEI/10.13039/501100011033.

Abstract

Although Antonio Batres Jáuregui was a leading scholar during the normativist period of Central American lexicography, his work has not received as much attention as that of others such as Gagini or Membreño. This paper analyzes the incidence of indigenous words collected in Batres's *Vicios del lenguaje y provincialismos de Guatemala* (Guatemala City: Encuadernación y Tipografía Nacional, 1892). It examines the relative frequency of these words in the text, their etymology, their lexical fields, and the text's explicit judgments concerning their use in order to better understand Batres Jáuregui's ideas and the evaluation of indigenous languages by nineteenth-century Central American lexicographers.

KEYWORDS: history of Latin American Spanish, lexicography, nineteenth century, Guatemala, indigenism.

FECHA DE RECEPCIÓN: 26/10/2022

FECHA DE ACEPTACIÓN: 11/01/2023



1. El siglo XIX y la historia léxica del español de Guatemala

A pesar de que es aún mucho lo que queda por analizar al respecto, no cabe duda de que en los últimos años el estudio histórico del vocabulario que se emplea en el español de Guatemala ha despertado un creciente interés entre los investigadores, lo que ha dado como resultado la aparición de múltiples investigaciones que permiten esbozar —ciertamente, aún de modo parcial y fragmentario— el desarrollo diacrónico que experimenta este nivel lingüístico en el país y que termina por configurar la importante personalidad léxica que hoy lo caracteriza. Aunque no es ahora el momento de hacer una revisión completa de tales estudios, es importante mencionar las relativamente abundantes aproximaciones dedicadas al siglo XVIII —a cuestiones como los indigenismos (Polo Cano, 2005; Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017) o la terminología propia de diversos ámbitos de especialidad como la política (Dym y Herrera Mena, 2014), el mestizaje (Ramírez Luengo, 2019), la medicina (Montero Lazcano, 2020; Ramírez Luengo, 2020, en prensa, en prensa b) o los tejidos (Martínez Puga, 2021)—, así como otras, bastante más limitadas en número y alcance, que se centran en

los primeros siglos de presencia española en la región (Ramírez Luengo y San Martín Gómez, 2020; Herrera Peña, 2021) o se dedican a analizar cuestiones concretas del siglo de las Independencias (Herrera Peña, 2021b; Ramírez Luengo, 2021, 2021b, en prensa c).

Por lo que se refiere a este último periodo histórico, la escasez de trabajos que se acaba de mencionar no sorprende, pues enlaza con el tradicional abandono que en general ha padecido el siglo XIX en los estudios acerca de la historia del español (Ramírez Luengo, 2012, p. 7; Buzek y Šincová, 2014, pp. 7-8; Company, 2017, p. 76). Sin embargo, conviene señalar que resulta especialmente grave en el caso de América si se tiene en cuenta la especial relevancia que, desde numerosos puntos de vista, posee este momento para la historia lingüística de la región: en efecto, fenómenos propios del Ochocientos como el fraccionamiento político de los antiguos virreinos en una serie de repúblicas independientes que favorecen la normativización de sus usos propios como estándar culto regional o la hispanización masiva de la población autóctona y la aparición/generalización de nuevas variedades de *español indigenizado* (Ramírez Luengo, 2011, pp. 15-17) demuestran claramente la indudable trascendencia de esta época para una mejor comprensión del devenir del español americano, y justifican sobradamente, por tanto, la necesidad de desarrollar nuevos y más detallados estudios que permitan conocer con mayor profundidad la situación que los diversos niveles del español americano —guatemalteco en este caso—, y muy especialmente su léxico, presentan en el corte cronológico mencionado.

2. El trabajo: objetivos, corpus y principios metodológicos

En este sentido, el presente trabajo tiene como propósito cubrir parcialmente los vacíos que se acaban de señalar en el párrafo anterior por medio del análisis lexicológico de los indigenismos que aparecen en los *Vicios del lenguaje y provincialismos de Guatemala* (Ciudad de Guatemala: Encuadernación y Tipografía

Nacional, 1892) de Antonio Batres Jáuregui, con la finalidad de conocer la incidencia que tienen estas voces en la que sin duda constituye la obra lexicográfica más importante, por su amplitud y por su calidad, de los publicados en este país en todo el siglo XIX. Partiendo, pues, de este propósito fundamental, los objetivos específicos que se persiguen son los siguientes: a) detectar los indigenismos léxicos recogidos en sus páginas y calibrar el peso de tales vocablos en el total de elementos censurados; b) llevar a cabo el análisis de estas voces según factores variados, tales como su origen etimológico, el campo léxico al que pertenecen o su carácter de americanismo desde una perspectiva de uso y dinámica de ese concepto (Ramírez Luengo, 2017);¹ c) indicar primeras dataciones, sean absolutas o centroamericanas, de determinados vocablos en las páginas de Batres; y en definitiva, d) comprender de manera más profunda —y a partir de los juicios de valor explícitos que, respecto a estas unidades léxicas, aparecen en el volumen analizado— las ideas lingüísticas que sustentan la reflexión de Batres Jáuregui sobre las lenguas autóctonas y, de forma más amplia, la valoración de estas en la lexicografía centroamericana de finales del siglo XIX.

Por lo que se refiere al autor, es bien sabido que Antonio Batres Jáuregui (Ciudad de Guatemala, 1847-1929) constituye uno de los intelectuales centroamericanos más importantes del siglo XIX, así como el filólogo más relevante de Guatemala en esta centuria. Nacido en el seno de una familia de la élite capitalina, estudia filosofía y derecho en la Universidad de San Carlos y desarrolla una importante carrera como político, diplomático y abogado, además de impartir docencia en esta misma casa de estudios; su interés y su cultivo de la historia y la filología lo llevan a pertenecer a instituciones culturales nacionales y extranjeras

¹ Como se dirá más adelante, la perspectiva de análisis que se va a utilizar diferencia claramente “entre *americanismo* —vocablo propio de América de acuerdo con su empleo en un momento concreto— e *indigenismo*, concepto que guarda relación con el origen etimológico de la voz en cuestión” (Ramírez Luengo, 2021b), lo que supone que ambos conceptos puedan coincidir en la misma unidad léxica, pero que esto no sea estrictamente necesario.

muy variadas como la Sociedad de Amigos del País, la Sociedad de Historia Diplomática de París, la Sociedad Literaria Hispano-Americana de Nueva York o la Academia de Ciencias de El Salvador, así como a fundar en 1887 la Academia Guatemalteca de la Lengua, que llega a dirigir durante cierto tiempo.² Por lo que se refiere a su pensamiento, la misma *BVFE* (2022) señala que “sus planteamientos, centrados en la figura del indígena, muestran tintes del racismo de la época y entran de lleno en el conflicto decimonónico de ‘civilización y barbarie’, planteando que el indígena debe asimilarse para lograr la ‘uniformidad biológica de la nación’”, en la línea de lo que constituye la ideología liberal (centro)americana propia de su época.³

Como era de esperarse, las creencias que se acaban de señalar terminan por reflejarse en su obra, e inevitablemente hacen acto de presencia en sus trabajos de corte filológico, fundamentalmente en sus *Vicios del lenguaje y provincialismos de Guatemala* (1892) y *El castellano en América* (1904), que en palabras de Quesada Pacheco (2008, p. 150) reflejan —junto a otros estudios de varios autores del Istmo— “una tendencia más bien de corte conservador, cuyo interés primordial era mantener la lengua y la literatura unidas”, resultado de “cierto temor ante la idea de una desarticulación de la lengua española en el Nuevo Mundo”. Este hecho determina que los estudiosos centroamericanos del siglo XIX se entreguen a la tarea de “escribir gramáticas y diccionarios que condenaran todo tipo de expresión dialectal que atentara contra la unidad lingüística” (Quesada Pacheco, 2008, p. 151),⁴ en una tarea que resulta especialmente clara en

² Toda esta información está tomada de la ficha biobibliográfica que la Biblioteca Virtual de la Filología Española (*BVFE*) dedica al autor, en concreto en <www.bvfe.es/es/autor/9310-batres-jauregui-antonio.html>. Para una revisión más detallada de su vida y de su obra como historiador y filólogo, véase también Erickson (1942).

³ Un detallado análisis de los postulados políticos e ideológicos de Batres Jáuregui y de su relación con el liberalismo centroamericano decimonónico se encuentra en Casaus Arzú (1999) y García Giráldez (2009).

⁴ Posteriormente se verá, sin embargo, que esta condena no es tan radical —al menos en el caso de Batres, y en lo que se refiere a los indigenismos—, lo que obligará a matizar parcialmente esta visión.

una obra como la que se va a analizar en estas páginas, cuya denominación como *vicios del lenguaje* y *provincialismos* supone, en realidad, toda una declaración de sus intenciones: presentar —a partir del modelo de ejemplaridad que supone la norma peninsular española— “aquellas voces mal pronunciadas junto con frases frecuentemente incorrectas que, tal como indica el título, suponen un vicio que afea el lenguaje” (BVFE, 2022).⁵

Teniendo en cuenta lo que se acaba de señalar, cabe preguntarse cuál puede ser el interés lexicológico de una obra como los *Vicios del lenguaje* de Batres Jáuregui, y a este respecto es importante hacer hincapié en tres cuestiones de muy diversa índole: por un lado, en la inexistencia hasta el momento de corpus filológicamente fiables del ochocientos guatemalteco que permitan analizar el léxico empleado en esta variedad diatópica en tal momento de su historia; por otro, en la riqueza y amplitud de los fondos que atesora este diccionario, cuyas más de 500 páginas y 1800 entradas —entre voces y locuciones (Quesada Pacheco, 2003, p. 552)— lo transforman en una de las mejores y más ricas fuentes para el análisis de la cuestión; por último, en la propia finalidad prescriptiva de la obra, que necesariamente determina la aparición en ella de vocablos de uso frecuente y cotidiano en la Guatemala de la época,⁶ así como la plasmación en sus páginas de informaciones de tipo sociolingüístico y diafásico sobre determinados vocablos que reflejan lo que se puede denominar la *mirada del filólogo* (Ramírez Luengo, en prensa d) y que resultan fundamentales para

⁵ Para un análisis de las ideologías que articulan y dan sentido a la obra filológica de Batres Jáuregui —cuestión de indudable relevancia para su mejor comprensión, pero que queda fuera de los intereses lexicológicos de estas páginas—, resulta del todo fundamental la consulta del muy reciente estudio de Rivas Hidalgo (2022).

⁶ Dado que, evidentemente, no tiene sentido que un purista como Batres Jáuregui utilice las páginas de su texto para corregir vocablos que no se emplean en el país y que, por tanto, no suponen un ataque a esa unidad de la lengua española que se procura (Quesada Pacheco, 2003, p. 549); en este sentido, no cabe duda de que esta obra —y en general todas las pertenecientes al periodo que el ya mencionado Quesada Pacheco (2003, p. 555) denomina *etapa purista* de la lexicografía del español de América Central— constituyen en realidad auténticos tratados de dialectología, cuyo análisis desde esta perspectiva puede aportar datos de notable importancia para la reconstrucción histórica del español centroamericano.

obtener una valoración más precisa y más completa de las voces que se emplean en Guatemala en los últimos años del siglo XIX.⁷ De este modo, no cabe duda de que la conjunción de factores tan diversos —pero al mismo tiempo tan relevantes— como los que se acaban de mencionar termina por transformar el texto lexicográfico de Batres Jáuregui en un documento de primera importancia a la hora de estudiar el vocabulario decimonónico de esta variedad diatópica, y de ahí que sea preciso prestarle más atención de la que por el momento le han dedicado en general los investigadores interesados en la historia léxica del español americano.

Por otro lado, una mejor comprensión del estudio que se desarrollará a continuación exige explicitar los principios metodológicos que se han aplicado para dar respuesta a los objetivos propuestos. A este respecto, parece pertinente mencionar que estas páginas siguen el esquema de trabajo que, para el estudio del indigenismo en el español centroamericano de los siglos XVIII y XIX, se ha empleado ya con anterioridad —y con resultados satisfactorios— en Ramírez Luengo (2019b), tanto en lo que se refiere a la organización de los aspectos que interesa analizar como también en las estrategias empleadas para ello; así mismo, se retoma también en estas páginas la interpretación del *americanismo* del estudio mencionado, que define este concepto como aquella unidad léxica identificadora del “habla urbana, popular o culta, o ambas, de América y cuyo uso muy frecuente y cotidiano distancia la variedad americana respecto del español peninsular” (Ramírez Luengo, 2015, p. 116), algo que necesaria-

⁷ En este sentido, son especialmente interesantes las abundantes notas de Batres Jáuregui sobre la valoración social de determinados vocablos y fenómenos en su época —a manera de ejemplo, define *charás* como “interjección muy baja y vulgar, que solo se escucha entre la hez del pueblo” (p. 205), indica que en Guatemala “ha quedado el uso de *masque* relegado a la gente baja” (p. 380) y que “no es de la gente zafia eso de llamar *policía* al agente de orden público” (p. 551), o señala el mantenimiento popular de formas antiguas como *vide/vido* (p. 551) y de la aspiración procedente de la /f/ inicial en vocablos concretos como *jalar*, *jacha*, *jarto* y *jeder* (p. 343)—, pues tales informaciones permiten comprender mejor la variación inherente a un estado de lengua y aportan, así, “datos de interés sobre la distribución y el uso de determinados elementos cuya consecución por otros medios resulta notablemente compleja” (Ramírez Luengo, en prensa d).

mente implica reconocer la existencia de tres subtipos de americanismos (*puros*, *semánticos* y *de frecuencia*)⁸ y aceptar el carácter dinámico de los mismos, dado que la interpretación de un vocablo como tal “no se mantiene inalterada a través del tiempo, sino que puede variar a lo largo de la historia, dependiendo de los procesos de expansión o reducción geográfica que experimenten las diferentes unidades léxicas” (Ramírez Luengo, 2012b, p. 398).

3. El indigenismo en las ideas lingüísticas de Batres Jáuregui⁹

No cabe duda de que un título como *Vicios del lenguaje y provincialismos de Guatemala* necesariamente parece identificar a su

⁸ Por cuanto no todos los vocablos responden a la definición facilitada de la misma manera. De este modo, mientras que los *americanismos puros* se interpretan como aquellas ‘voces empleadas en el español general de América inexistentes en el español peninsular general’, los *semánticos* se corresponden con las ‘voces y construcciones formalmente compartidas con el español peninsular, pero que han desarrollado en América valores semánticos propios’ y el subtipo *de frecuencia* agrupa ‘voces o construcciones compartidos, en forma y significado, con el español peninsular castellano, pero que muestran en América una mucho mayor frecuencia de empleo y de generalización’ (Company, 2010, p. XVII).

⁹ Por supuesto, la importancia del autor en la historiografía lingüística centroamericana y la riqueza de información que aportan a este respecto las páginas de los *Vicios del lenguaje* hacen del todo necesario continuar la labor de Rivas Hidalgo (2022) y profundizar en esta cuestión, que se debería completar además con un análisis de las abundantísimas fuentes que Batres Jáuregui emplea en el texto para justificar sus juicios, tanto literarias —con clásicos hispánicos (*Mío Cid*, *Lazarillo*, Santa Teresa, Ercilla, Quevedo, Góngora, Cervantes, Fuentes y Guzmán, Jovellanos, etcétera) pero también escritores contemporáneos, españoles y americanos (Núñez de Arce, Bécquer, Zorrilla, Menéndez Pelayo; Goyena, Palma, Isaacs, Salomé Jil, Rivera Maestre)— como lingüísticas (entre otras, Nebrija, Terreros, Bello, Salvá, Pichardo, Arona, Cuervo, Rodríguez o Irisarri, además de las obras académicas); este extenso listado demuestra no solo la amplia cultura del lexicógrafo guatemalteco y su profundo conocimiento de los estudios filológicos de la época —habida cuenta de la aparición en su diccionario de los nombres fundamentales del momento—, sino también que en modo alguno se lo puede considerar un simple aficionado o diletante. Desgraciadamente, tales aspectos superan en mucho los objetivos del presente estudio, de manera que este apartado pretende ser solo un primer acercamiento a la temática señalada que haga hincapié —al tiempo que demuestra el interés de acometer su estudio— en aquellos aspectos que, desde este punto de vista, guardan una relación más estrecha con el indigenismo.

autor con esa corriente purista que, para la lexicografía centroamericana, establece Quesada Pacheco (2003, p. 550) en el siglo XIX, y de hecho son muchas las entradas del texto que parecen avalar esta interpretación, habida cuenta de la denuncia que Batres Jáuregui hace de lo que considera usos *incorrectos* en la pronunciación —sean fenómenos ya asentados en el país, como el yeísmo (p. 552), o sean vocablos concretos (*aniegan, bravian-do, espelma, espeutro, gomitár, sectiembre*)—, en la morfología —los géneros discordantes con la norma de *reúma* y *sartén*, la confusión de prefijos *anti-/ante-* (*antidiluviano*), ciertas formas verbales vulgares y/o analógicas (*haiga, quedarán; andó, conducí, desdecí, maldecí*)— o en la sintaxis, donde critica la pluralización del *haber* existencial (pp. 316, 335), el denominado *hasta mexicano* (p. 319), el uso transitivo de *arder* (p. 108) y muy especialmente el voseo (pp. 241, 247, 319, 392, 548). Por supuesto, su interés alcanza también al léxico, donde el intelectual guatemalteco no solo señala modificaciones semánticas americanas (*apearse, alborotos, apercibirse, bandada*) y discute la necesidad de incorporar ciertos neologismos (*locomotiva, presupuestar, provisorio, tramitar, valorizar*), sino que además critica el empleo de arcaísmos (*aprevenido, atufado, chinchón, escurana*), anglicismos (*brequé, budín, espiche, lunch/lonche*) y muy especialmente galicismos puros o semánticos (*acusar, bouquet, etiqueta, gros, ilusionarse, petipieza, pluscafé*), en línea, pues, con lo que suelen ser las tendencias habituales en las obras de naturaleza purista que se publican en su época.

Como era de esperar, también los indigenismos hacen acto de presencia en estos *Vicios del lenguaje y provincialismos de Guatemala*, si bien es importante mencionar que la actitud que Batres Jáuregui muestra hacia ellos está muy lejos de la denuncia y la crítica que —respecto a otros elementos— se ha mencionado en el párrafo anterior: así, tras recordar que “los numerosos idiomas indígenas prestaronle [= al español americano] rico contingente, al punto que, mientras viva, guardará restos del quiché, del mexicano, del quichua, del guaraní y de todas las principales lenguas que aquí hablaban al llegar los capitanes iberos” (p. 30), el autor indica una serie de factores que, a su

entender, justifican la necesidad de incorporar los vocablos de origen amerindio al léxico español, entre los que se pueden destacar su uso generalizado entre los hablantes del continente (“las apadrina el uso de millares de hombres”; p. 32),¹⁰ su idoneidad para expresar conceptos y referentes novedosos (“se introducen justificados por la necesidad de denominar objetos o seres nuevos”; p. 28) o la relevancia que poseen para la propia identidad lingüística de los hispanoamericanos (“están en la condición modesta de provincialismos nuestros; pero que para nosotros tienen la importancia que en la familia se atribuye a las reliquias abolengas”; pp. 31-32). Salta a la vista, por tanto, que para Baires Jáuregui el indigenismo, lejos de ser un problema, supone en realidad un enriquecimiento del español (“el rico idioma de Don Quijote y Sancho habría que hallar con el tiempo, inagotable venero de elementos lexicográficos, dignos de tomarse en cuenta [...], que lejos de amenguar el habla castellana, dale más valor, riqueza y gallardía”; pp. 32-33), el cual en modo alguno puede circunscribirse a los estrechos límites de sus orígenes latinos o de la norma peninsular:

Aumentóse asombrosamente el vocabulario usual, con voces autóctonas, aplicadas a objetos y usos distintos a los de España; voces que son americanismos de simpático sonido y regular estructura, dignos de figurar en el diccionario de la lengua, ¿ni qué más tiene que una voz descienda del latín o del árabe, o se derive del quichua o el cackchiquel, si se emplea por una colectividad respetable, de los veinticinco millones de hombres que hablan español en este continente? Nadie ha pretendido jamás que sólo el lenguaje que se oye al borde del Manzanares o las palabras que se escuchan en la calle de Valverde sean las que registre el léxico de la lengua (p. 33).

¹⁰ De hecho, esta idea del uso —según el ilustre guatemalteco, “árbitro del lenguaje” (p. 366)— aparece de manera constante en el texto, y se refleja, entre otras cuestiones, en su reiterado interés por señalar la distribución diatópica de determinados elementos y su presencia en regiones americanas tan variadas como México, Perú, El Salvador, Colombia, Chile o Argentina (pp. 246, 258, 266, 323, 352, etcétera).

Así las cosas —y a la luz de lo que se acaba de señalar—, no puede sorprender que las entradas que se dedican a las voces de origen amerindio en los *Vicios del lenguaje y provincialismos de Guatemala* se caractericen en general por constituir una explicación —en numerosas ocasiones de naturaleza enciclopédica—¹¹ sobre el término en cuestión, en la que se precisa su origen etimológico, su distribución diatópica y su significado preciso, muchas veces en claro contraste con la información que facilita al respecto el *Diccionario* de la Real Academia Española, tal y como demuestran los ejemplos siguientes (ejemplos 1-3):

- (1) En Guatemala significa una escalerilla de tablas, que sirve para llevar algo a cuestras (...). Es palabra mexicana, adoptada como tal por el diccionario de la Academia Española, que trae *cacaxtle* (p. 151).
- (2) Así llamamos a lo que generalmente en América llaman *chacra*, que según el Diccionario de la Academia, es una vivienda rústica y aislada; pero con perdón del ilustre Cuerpo, *chacra* lo que quiere decir es una propiedad rústica pequeña, sembrada de árboles frutales (p. 195).
- (3) esta palabra se deriva del mexicano *exotl*, que significa la vaina del frijol, cuando está verde. Como provincialismo de México figura aquel vocablo en el Diccionario de la Academia. Conste, pues, que también es voz que se usa en Guatemala (p. 264).

Se puede concluir, por tanto, que la presencia de indigenismos en la obra que se está analizando no se debe interpretar como una sanción contra estas voces, sino que, muy al contrario, supone una reivindicación de lo americano que implícitamente tiene como objetivo señalar la necesidad de incorporar tales vocablos a la obra académica, estableciendo para ello un diálogo

¹¹ Algo especialmente notorio en entradas específicas como *copalchí* (pp. 185-186), *chumpipe* (pp. 233-234), *guarumo* (pp. 311-312) o, por supuesto, *quetzal* (pp. 478-482), ave nacional y emblema de Guatemala, en cuya extensísima explicación se hace muy evidente, por razones obvias, la subjetividad del autor.

entre Batres Jáuregui y la Docta Casa que también se produce en el caso de otros coetáneos como el peruano Ricardo Palma, y que sin duda obliga a considerar al autor —tanto por la atención prestada a su propia variedad diatópica como por el valor que le concede— un ejemplo más del paso de lo puramente preceptivo a lo descriptivo que se detecta a finales del siglo XIX en la lexicografía centroamericana (Quesada Pacheco, 2003, p. 550).

4. Batres Jáuregui y la historia del léxico guatemalteco: los indigenismos

A partir de todo lo que se ha señalado en el punto anterior, no puede sorprender que la presencia de indigenismos en la obra que se está analizando resulte muy abundante, y que se descubran tanto en el leuario como en el mismo cuerpo de las entradas, donde en ocasiones se emplean para definir o explicar los vocablos que interesa comentar (ejemplos 4-5).

- (4) Es provincial la acepción de *berrinche* cuando los soldados dan ese nombre a una comida hecha de tortilla o *totoposte* con agua y sal (p. 134).
- (5) *Comal* o *cumal* es un disco grande, delgado, de barro, con bordes, en el que los indios cocían y aún cuecen el *tazcal* o torta de maíz (p. 182).

Tomando ahora en cuenta únicamente las voces que aparecen lematizadas, conviene señalar en primer lugar que los vocablos de indudable origen amerindio ascienden a un total de 230 —en concreto 197 simples y 33 derivados—, a los que probablemente se puedan añadir otros que no se recogen en el corpus lexicográfico consultado¹² o no presentan en él su etimología, pero

¹² En concreto, Santamaría (1959), el *Diccionario Crítico Etimológico Castellano e Hispánico* (DCECH, 1980-91), Morinigo (1998), el *Diccionario de Americanismos* (DAMER, 2010) y el *Diccionario de la Lengua Española* (DLE, 2014), así como —más puntualmente— los estudios de Buesa Oliver y Enguita Utrilla (1992), DHECR (1995)

cuya forma parece denotar esta misma procedencia (entre otros, *chojín*, *chutuy*, *loroco*, *güiriche*, *güisespín*, etc.).¹³ De este modo, si se considera que los *Vicios del lenguaje* cuentan con 1800 entradas (Quesada Pacheco, 2003, p. 550), salta a la vista que los indigenismos equivalen al menos al 12.77% del total de sus entradas, un porcentaje sin duda muy destacable que demuestra el interés que, desde el punto de vista cuantitativo, posee la obra analizada para la historia léxica del español de Guatemala, y más concretamente para el estudio de los préstamos amerindios, de aparición no tan abundante en otras tipologías textuales.¹⁴

Dicho esto, conviene aportar ahora los 230 vocablos de indudable origen indígena que aparecen lematizados en la obra de Batres Jáuregui, y que son los que aparecen en el listado siguiente: por lo que se refiere a las voces simples, *aguacate*, *agüizote*, *aiguaixte*, *amol*, *anona*, *apasote*, *apaste*, *atol*, *ayote*, *baja-reque*, *bejuco*, *bucul*, *cacao*, *cacaxte*, *caco* ‘hicaco’, *caimito*, *caite*,

y Herranz (2017). En ocasiones se ha recurrido también a las informaciones que aporta sobre esta cuestión el propio Batres Jáuregui, muy interesado en la cuestión etimológica —que aparece profusamente en citas del estilo de “esta palabra tan usada entre nosotros es cakchiquel, y denota la camisa de las indias” (*huiopil*; p. 329), “es nombre quechua, y significa entre nosotros y en otras regiones de la América la sabrosa fruta que los botánicos conocen con el mismo nombre” (*papaya*; p. 421), “al cuesco del zapote llámanle sapuyulo, que sospechamos sea voz indígena” (*sapuyulo*; p. 501)—, si bien es importante mencionar que algunos de sus juicios no carecen de errores: “este nombre viene, según la Academia, del haitiano *hacamayo*; pero nos inclinamos a creer que se deriva más bien de *guaca*, que quiere decir *sagrada* y *maya*, tribu de indios que se hallaba al norte de México” (*guacamaya*; p. 306); “se deriva del cumanagoto *huarapu*, que quiere decir *tasa*” (*guarapo*; p. 312).

¹³ También quedan fuera de este listado algunos vocablos de claro origen amerindio que se pueden englobar dentro del amplio campo de la onomástica (*chontal*, *jicaque*, *jipijapa*, *pirú*), dado que su análisis exige un acercamiento parcialmente diferente al de los nombres comunes; con todo, tanto esta categoría como las voces de etimología dudosa permiten postular una presencia de términos tomados de las lenguas autóctonas que va más allá de los 230 que se acaba de mencionar.

¹⁴ A este respecto —y a pesar de que las discrepancias cronológicas dificultan la comparación—, recuérdese por ejemplo que en un conjunto de 79 documentos de archivo generados en Guatemala durante el siglo XVIII se detectan apenas 18 indigenismos (Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, pp. 113-114), número notablemente más parco del que ofrecen las páginas de Batres Jáuregui.

cajete, calpul, calzonte, camagua/camagüe, cancha, canchalagua, canche, cantil, cempoasúchil, censonte, chácara, chacha, chachaguante, chajal, chalchigüite, chan, chancaca, chaneque, chapulín, chapupo, chara, chay, chayote, checa, chele, chibola, chicalote, chicha, chiche, chichicaste, chichigua, chigua, chilacayote, chilaquila, chilate, chilca, chile, chiltepe, chiltote, china, chinama, chingaste, chipe, chipilín, chipote, chipuste, chiquigüite, chiquirín, chirmol, chocoyo, chongo, chuco, cipote, claco, colochó, comal, comején, copalchí, copinol, cotuza, coyol, coyote, dividivi, ejote, eloatol, elote, guaca, guacal, guacamaya, guachipilín, guaco, guaguá, guaje, guajiro, guanaba, guanaco, guapinol, guarapo, guaro, guarumo, guasanga, guate, guayaba, güegüecho, güicoy, güisquil, huarahua, huipil, huisache, hule, izote, izquisuchil, jaba, jícama, jícara, jiote, jiquilite, jobo, jocote, jolote, juilín, jute, lora, manía ‘mani’, mapache, matate, mazacuata, maztlate, mecapal, mecate, mica, milpa, miltomate, mole, nacascolote, nagua, nance, nigua, nopal, ocote, pacaya, pachte, papa, papalote, papaya, pepesca, petate, pijije, pilguanejo, pilixte, piloy, pinol, pita, pitahaya, pizote, puchito, pulique, pupusa, pusunque, quequexque, quetzal, quijinicuil, sanate, sapuyulo, siguapate, sucucho, tabanco, tacuazín, taltusa, tamal, tamehua, tanate, tapalcate, tapalcúa, tapaljocote, tapesco, tayuyo, tazol, tecolote, tecomate, temascal, tenamaste, tepemechín, tepescuinte, tepocate, tequio, tetunte, tihuilote, tiste, tizate, tol, totopoxte, tul, tuna, tusa, yagual, yuca, zacate/zacatón, zopilote y zute; en cuanto a las derivadas, achimero, achucuyado, aguacatal, atolillo, azopilatado, cacaguatal, camotillo, chichona, chilar, chilero, chinear, chipear, chipioso, chiviviscal, chirmoloso, chojolería, cholojera, chumpipada, chumpipear, empetatar, enchichicastado, enchiladas, enzacatarse, guacaluda, guanacada, guatal, guayabal, micada, taltusero, tamalera, tomatal, totopón y yucal.

Dentro del análisis de las palabras simples —a las que se va a circunscribir el presente estudio—,¹⁵ resulta de interés cen-

¹⁵ Esta reducción del análisis a las voces simples se justifica por el carácter propiamente hispánico de los vocablos derivados, que son el resultado de un proceso morfológico de la lengua española por más que el elemento que les sirve de base sea de

trarse en primer lugar en determinar cuál es su origen etimológico, pues esta cuestión permite obtener una primera idea sobre las lenguas que de manera más abundante contribuyen a enriquecer el léxico del español de Guatemala y, por extensión, de Centroamérica. De este modo, la revisión de la información que respecto a este asunto ofrece el corpus lexicográfico de referencia ya mencionado anteriormente permite establecer la tabla siguiente (tabla 1):

LENGUA	CASOS	VOCES
Náhuatl	142 (72.08%)	<i>aguacate, agüüzote, aiguaihte, amol, apasote, apaste, atol, ayote, cacao, cacaxte, caite, cajete, calpul, calzonte, camagua/camagüe, cempoasúchil, censonte, chacha, chachaguante, chalchigüite, chan, chaneque, chapulín, chayote, chicalote, chiche, chichicaste, chichigua, chigua, chilacayote, chilaquila, chilate, chile, chiltepe, chiltote, chinama, chingaste, chipe, chipilín, chipote, chipuste, chiquigüite, chiquirín, chirmol, chocoyo, chongo, chuco, cipote, claco, colochó, comal, copalchí, copinol, cotuza, coyol, coyote, ejote, eloatol, elote, guacal, guachipilín, guaje, guapinol, guate, güegüecho, güisquil, huipil, huisache, hule, izote, izquisuchil, jícama, jícara, jiote, jiquilite, jocote, jolote, juilín, jute, mapache, matate, mazacuata, maztlate, mecapan, mecate, milpa, miltomate, mole, nacascalote, nance, nopal, ocote, pachte, papalote, pepesca, petate, pijije, pilguanejo, pilixte, piloy, pinol, pizote, pupusa, quequexque, quetzal, sanate, sapuyulo, siguapate, tabanco, tacuazín, taltusa, tamal, tanate, tapalcate, tapaljocote, tapesco, tazol, tecolote, tecomate, temascal, tenamaste, tepemechín, tepescuinte, tepocate, tequio, tetunte, tiste, tizate, totopoxte, tul, tusa, yagual, zacate/zacatón, zopilote, chapupo, chele, güicoy, pusunque, quijinicuil, tapalcúa, tihuilote, pulique</i>

naturaleza amerindia; además, el hecho de que algunos de estos elementos (*chumpipada, chumpipear, guanacada*) cuenten en el listado con el indigenismo que les sirve de base (*chumpipe, guanaco*) parece aconsejar dejarlos fuera de los análisis porcentuales, cuyos resultados podrían verse alterados por estas duplicaciones.

Leng. antill.	22 (11.16%)	<i>anona, bajareque, bejuco, caco</i> ‘hicaco’, <i>caimito, comején, guacamaya, guaco, guajiro, guanaba, guarumo, guayaba, jaba, jobo, manía</i> ‘maní’, <i>nagua, nigua, pita, pitahaya, tayuyo, tuna, yuca</i>
Quechua	15 (7.61%)	<i>cancha, chácara, chilca, china, guaca, guaguá, guanaco, guarapo, guaro, pacaya, papa, puchito, sucucho, chancaca, huarahua</i>
Leng. mayas	9 (4.56%)	<i>canche, cantil, chibola, tol, tamehua, bucul, chajal, chay, zute</i>
Caribe	4 (2.03%)	<i>guasanga, lora, mica, papaya</i>
Cuna	1 (0.50%)	<i>chicha</i>
Mapudungun	1 (0.50%)	<i>canchalagua</i>
Sin etimol.	3 (1.52%)	<i>chara, checa, dividivi</i>
TOTAL	197 (100%)	

Tabla 1. Origen etimológico de los indigenismos del corpus

Dejando aparte tres vocablos cuya etimología no se ha podido obtener (*chara, checa, dividivi*), salta a la vista que es el náhuatl la lengua que resulta claramente predominante en los *Vicios del lenguaje*, con 142 elementos que equivalen a casi el 75% de todos los indigenismos del corpus, seguido a gran distancia por las lenguas antillanas y el quechua, que presentan unos porcentajes muy inferiores, del 11% y el 7% respectivamente; más allá de estos idiomas, los aportes del resto no pasan de ser meramente testimoniales, y en todos los casos presentan valores inferiores al 5%: nueve ejemplos tomados de las lenguas mayas¹⁶ y cuatro préstamos caribes, así como un único vocablo procedente del cuna y otro del mapudungun. Se esboza, así, una situación que concuerda muy bien con lo que se podía esperar, dado que refleja de manera muy clara tanto lo que numerosos investigadores han señalado reiteradamente sobre la influencia

¹⁶ Como es fácil de suponer, se utiliza en este caso el sintagma plural *lenguas mayas* porque tales préstamos parecen proceder —al menos, a la luz de lo mencionado por el propio Batres Jáuregui— de idiomas distintos, tales como el maya yucateco (*canche*), el cakchiquel (*tamehua*) y muy especialmente el quiché (*bucul, chajal, chay, zute*).

léxica de los idiomas amerindios en el español americano —esto es, la importancia del náhuatl en los territorios del virreinato novohispano, la escasa relevancia cuantitativa del aporte maya o la extensión de los antillanismos más allá de su solar originario (Ramírez Luengo, 2007, pp. 76-9)—, como lo que constatan otros estudios sobre la historia léxica de Centroamérica y/o Guatemala (Polo Cano, 2005, pp. 191-192; Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 115; Ramírez Luengo, 2019b, pp. 255-256, 2021b), en una muestra más de la continuidad fundamental que, con respecto a las tendencias de uso de estos elementos, se descubre en estas variedades diatópicas del español con independencia de los momentos históricos o de los corpus considerados.

Al mismo tiempo, la revisión de los datos etimológicos que aporta el texto de Batres Jáuregui permite constatar otros aspectos de interés, que se pueden cifrar fundamentalmente en tres: en primer lugar, es importante hacer hincapié en la cantidad de sistemas lingüísticos que, de acuerdo con el corpus aportan vocablos al español guatemalteco, así como su empleo en ámbitos geográficos muy distantes, pues ambas circunstancias evidencian muy a las claras la complejidad que supone la configuración léxica de las hablas americanas, resultante —entre otras cuestiones— no solo del contacto directo con los idiomas amerindios, sino también de los trasvases de vocabulario y de las influencias mutuas que establecen entre sí tales hablas a lo largo de la historia; en segundo lugar —y en clara relación con lo anterior—, resulta también relevante constatar la abundante presencia de quechuismos en esta obra, pues su empleo en Centroamérica durante el siglo XIX demuestra la necesidad de relativizar la frontera que tradicionalmente se establece en el sur de Colombia para la difusión de estos elementos; por último, es también muy valiosa la aparición en sus páginas de cierto número de mayismos, cuya escasez cuantitativa se ve compensada por su indudable importancia cualitativa a la hora de dotar de personalidad al español guatemalteco de la época, así como por la dificultad que supone la detección de estos elementos en otras tipologías textuales, cuestión que refuerza aún más la trascendencia que posee esta obra para el estudio histórico del español de Guatemala.

Por otro lado, el acercamiento a estos vocablos desde el punto de vista etimológico se puede complementar con un análisis de los campos léxicos a los que tales elementos se incorporan, pues este hecho “permite obtener una idea más precisa acerca de las realidades que de manera predominante son expresadas con préstamos de las lenguas autóctonas” (Ramírez Luengo 2019b, p. 256); en este sentido, la clasificación que —con base en la de Ramírez Luengo (2019b)— se ha aplicado a este corpus arroja los resultados siguientes (tabla 2):

CAMPO LÉXICO	CASOS	VOCES
Agricultura/ ganadería	41 (20.81%)	<i>aguacate, anona, apasote, ayote, cacao, caco</i> ‘hicaco’, <i>caimito, camagua/camagüe, chácara, chan, chayote, chilacayote, chile, chiltepe, chipilín, coyol, ejote, elote, guanaba, guate, guayaba, güicoy, güisquil, jícama, jocote, manía</i> ‘maní’, <i>milpa, miltomate, nance, nopal, papa, papaya, piloy, pitahaya, tamehua, tapaljacote, tazol, tuna, tusa, yuca, zacate/zacatón</i>
Enseres / utensilios	35 (17.76%)	<i>apaste, bucul, cacaxte, caite, cajete, calzonte, chachaguante, chalchigüite, chay, chinama, chiquigüite, comal, guacal, guaje, huipil, jaba, jícara, matate, maztlate, mecapan, mecate, nagua, papalote, petate, tabanco, tanate, tapalcate, tapesco, tecomate, tenamaste, tetunte, tizate, tol, yagual, zute</i>
Fauna	34 (17.25%)	<i>cantil, censonte, chacha, chapulín, chara, chiltote, chiquirín, chocoyo, comején, cotuza, coyote, guacamaya, guaro, jolote, juilín, jute, lora, mapache, mazacuata, mica, nigua, pepesca, pijije, pizote, quetzal, sanate, tacuazín, taltusa, tapalcúa, tecolote, tepemehcín, tepescuinte, tepocate, zopilote</i>
Flora	29 (14.72%)	<i>amol, bejuco, canchalagua, cempoasúchil, chicalote, chichicaste, chigua, chilca, copalchí, dividivi, guachipilín, guaco, guapinol, guarumo, izote, izquisuchil, jiquilite, jobo, nacascalote, ocote, pacaya, pachte, pita, quequexque, quijinicuil, sapuyulo, si-guapate, tihuilote, tul</i>

Alimentación	19 (9.64%)	<i>aiguaixte, atol, chancaca, checa, chicha, chilaquila, chilate, chirmol, elotol, guarapo, mole, pinol, pulique, pupusa, pusunque, tamal, tayuyo, tiste, totopoxte</i>
Partes del cuerpo/ características físicas	10 (5.07%)	<i>canche, chele, chiche, chipuste, chongo, cipote, colochó, güegüecho, huarahua, pilixte</i>
Sociedad/ organización social	10 (5.07%)	<i>calpul, chajal, chichigua, china, guaguá, guajiro, guanaco, guasanga, huisache, pilguanejo</i>
Industria/ construcción	9 (4.56%)	<i>bajareque, cancha, chapupo, claco, copinol, guaca, hule, sucucho, temascal</i>
Otros	10 (5.07%)	<i>agüizote, chaneque, chibola, chingaste, chipec, chipote, chuco, jote, puchito, tequío</i>
TOTAL	197 (100%)	

Tabla 2. Distribución en campos léxicos de los indigenismos del corpus

Así pues, quizá lo primero que se debe resaltar es la presencia de préstamos amerindios en múltiples campos léxicos de naturaleza enormemente variada —como la agricultura, la alimentación, los enseres y utensilios, la organización social o las partes del cuerpo y las características físicas—, pues este hecho constata una vez más que la incorporación de indigenismos, extendida por todo el vocabulario, supone una auténtica reorganización de este nivel lingüístico que resulta, además, fundamental para dotar de personalidad a las diferentes variedades americanas del español. Al mismo tiempo, es importante mencionar que esta generalización del indigenismo por múltiples ámbitos de la realidad no oculta la también evidente concentración mayoritaria en ciertas realidades concretas como la agricultura/ganadería, los enseres y utensilios, la fauna y la flora —que en conjunto representan el 70.55% de todos estos vocablos, equivalentes a 139 de los 197 localizados en el corpus—, en una situación que enlaza con la que se detecta en otros estudios sobre esta cuestión en Guatemala o en América en general (Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 117; Mejías, 1980, p. 19; Ramírez Luengo, 2007, pp. 74-75), y que se explica por “la

originalidad que muestra América en estos aspectos” y, como consecuencia de ello, por “la necesidad que tiene el español de dar nombre a unos referentes desconocidos que carecen de él en la lengua” (Ramírez Luengo, 2019b, p. 258).¹⁷

En otro orden de cosas, es importante mencionar que las unidades léxicas que se registran en estos *Vicios del lenguaje* tienen también una gran trascendencia desde el punto de vista cronológico: en efecto, a pesar de que la modernidad de la obra (1892) pueda llevar a pensar lo contrario, lo cierto es que las escasas investigaciones desarrolladas hasta el momento sobre la temática tratada en estas páginas y la general ausencia —o escasa representación, al menos— de Guatemala en los corpus diacrónicos del español determina que en numerosas ocasiones los ejemplos registrados supongan la primera datación, absoluta o centroamericana, del elemento en cuestión, una nueva atestiguación de vocablos poco documentados en el pasado o incluso la única referencia histórica de voces registradas únicamente en sincronía, todo lo cual transforma a este texto en una fuente de primera importancia para el estudio del indigenismo y, por extensión, para la más completa reconstrucción de la historia léxica de la variedad guatemalteca del español.

Así pues —y a la luz de la información que ofrecen los grandes repositorios digitales (*CORDE*, *CORDIAM*, *LEXHISP*) y ciertas obras lexicográficas con datos sobre el desarrollo diacrónico del español centroamericano y guatemalteco (*DCECH*, *DHECR*; Quesada Pacheco, 2008; Ramírez Luengo, 2021)—, parecen pertenecer al último grupo mencionado, por ejemplo,

¹⁷ Con todo, es importante mencionar que no es este el único factor que explica la incorporación de indigenismos, habida cuenta de su aparición para hacer referencia a partes del cuerpo y características físicas (*coloch*o ‘rizo de pelo’, *pilixte* ‘débil o enfermizo’; *DAMER*, 2010: s.v. *coloch*o, *pilixte*) o a conceptos más o menos abstractos (*chuco* ‘sucio, desaseado’, *puchito* ‘porción pequeña de algo’; *DAMER*, 2010: s.v. *chuco*, *puchito*) que, evidentemente, existen —y poseen una denominación— antes de que los españoles lleguen a América; se hace necesario, por tanto, seguir investigando cuáles son las causas que explican la apropiación de estos vocablos por parte de los españoles, así como las circunstancias que explican que en ocasiones se sustituyan con ellos los nombres hispánicos tradicionales.

*chigua, aiguaihte, camagua/camagüe, cantil, chaneque, chapupo, checa, chele,*¹⁸ *chibola, chipuste, copinol, cotuza, guanaco,*¹⁹ *jute, mazacuata, pijije, pilixte, piloy, quijinicuil, tamehua, tapalcúa, tapaljcote, tihuilote, zute*, mientras que la presencia en Batres Jáuregui de *calzonte, chajal, chocoyo* y *copalchí* viene a engrosar las muy escasas ocurrencias de tales unidades —en todos los casos, un máximo de cinco en los corpus consultados—²⁰ con que se contaba hasta el momento. Por lo que se refiere a las primeras dataciones, los casos documentados en este corpus suponen la primera aparición centroamericana de 33 indigenismos (en concreto, de *caco* ‘hicao’, *canchalagua, cempoasúchil, chancaca, chara, chiche, chilacayote, chilaquila, chilca, china, chipilín, claco, comején, dividivi, guaguá, guajiro, guanaba, huarahua, huisache, jícama, jolote, manía* ‘maní’, *mecapal, miltomate, mole, nigua, pilguanejo, puchito, siguapate, sucucho, tayuyo, tepocate* y *yagual*), así como el testimonio más antiguo de *bucul, canche, chinama, chipe, chipote, chiquirín, chuco, cipote, eloatol, guasanga, güicoy, izote, matate, pulique, pupusa, pusunque, sapuyulo, tazol, tenamaste, tetunte* y *tol*, hasta hoy no registrados en su práctica totalidad antes de mediados del siglo xx.²¹

Finalmente, si hasta el momento se han analizado estos vocablos en tanto que indigenismos, parece importante también estudiarlos ahora en su carácter de *americanismos*, es decir, no tanto por su origen etimológico como por su empleo diatópico (Ramírez Luengo, 2017, p. 609) y, por ello, por su capacidad de dotar de personalidad léxica a una variedad geográfica concreta,

¹⁸ Con el valor propiamente guatemalteco de ‘legaña’ (DAMER, 2010: s.v. *chele*), que es el que aparece en la obra de Batres Jáuregui.

¹⁹ Al igual que en el caso anterior, con la significación de ‘persona poco inteligente y simplona’ (DAMER, 2010: s.v. *guanaco*) que apunta el lexicógrafo guatemalteco.

²⁰ En concreto, se descubren dos casos de *chajal* en Ramírez Luengo (2021), cuatro de *copalchí* en CORDE, LEXHISP y DHECR y cinco de *chocoyo*, tanto en el repositorio académico como en el DHECR; *calzonte*, por su parte, parecía ser un hápax hasta el momento, al aparecer en una única ocasión en LEXHISP.

²¹ Casi siempre, cabe señalar, en los *Cuentos de barro* (1943) del salvadoreño Salarrué o en *Hombres de maíz* (1945), de Miguel Ángel Asturias.

en este caso la guatemalteca.²² De este modo, mientras que en la aproximación anterior todos los términos coincidían en su carácter de préstamo de las lenguas originarias de América, la situación que muestran desde este nuevo punto de vista es muy otra, pues las importantes diferencias que se detectan entre ellos los sitúan en las diversas categorías que engloba el concepto *americanismo*.

Por lo tanto, a pesar de que *CORDE* no permite describir en profundidad el empleo español y americano de todas estas voces durante la época de Batres Jáuregui,²³ la distribución diatópica que se detecta en este corpus parece demostrar el carácter de *americanismos puros* (AP) de, entre otras, *atol*, *calpul*, *chachaguate*, *chalchigüite*, *chapulín*, *chichigua*, *jiote*, *jiquilite*, *ocote*, *pacaya* o *tecolote*, así como que *aguacate*, *cacao*, *chácara*, *nagua*, *papaya* o *yuca* reflejan las características propias de los *americanismos de frecuencia* (AF).²⁴ Por su parte, los datos constatan también que a finales del siglo XIX *guanaco* —‘mamífero camélido propio de los Andes’ en España (*CORDE*) frente al valor guatemalteco de ‘provinciano’ que apunta Batres (pp. 308-309)—

²² Se sigue, por tanto, la interpretación de este concepto que se señalaba más arriba, y que supone la existencia de diversos tipos de americanismos (*puro*, *semántico*, *de frecuencia*) y el carácter dinámico del concepto; esta última idea obliga a circunscribir el análisis a una sincronía específica, que en este caso se ha cifrado en el periodo 1842-1942 —es decir, 50 años previos y 50 posteriores a la datación de los *Vicios del lenguaje*—, siguiendo para ello el ejemplo y las motivaciones ya expuestas en trabajos previos (Ramírez Luengo, 2021b).

²³ De hecho, un gran número de ellas —a manera de ejemplo, *chicalote*, *pinol* o *quequexque*— no aparecen en ninguna ocasión en el corpus de la RAE, si bien este hecho se puede interpretar como una prueba a favor de la interpretación como *americanismos puros*, algo que parece verse corroborado también por el desconocimiento actual que de tales vocablos existe en la variedad peninsular; en otras ocasiones los ejemplos registrados son notablemente escasos, por lo que su carácter probatorio no es tan alto como sería de desear.

²⁴ Dado que, aunque se localizan a ambos lados del Atlántico, presentan diferencias cuantitativas de peso en lo que se refiere a su aparición, tal y como se muestra a continuación en casos por millón de palabras (CMP): *aguacate* (3.93 CMP - Am. vs. 0.35 CMP - Esp.), *cacao* (13.04 CMP - Am. vs. 1.88 CMP - Esp.), *chácara* (3.02 CMP - Am. vs. 0.31 CMP - Esp.), *nagua* (1.32 CMP - Am. vs. 0.18 CMP - Esp.), *papaya* (1.84 CMP - Am. vs. 0.12 CMP - Esp.) o *yuca* (6.78 CMP - Am. vs. 0.27 CMP - Esp.).

se debe entender como *americanismo semántico* (AS), mientras que *jícara* y *nopal*, presentes por igual en España y América, se han transformado ya en estos momentos en voces generales (VG) —sin capacidad, pues, para identificar diatópicamente al español de Guatemala—, en un nuevo ejemplo de los procesos de *desamericanización* (Ramírez Luengo, 2017b, p. 103) que modifican dinámicamente la valoración de determinadas voces y que, poco atendidos aún por los investigadores, resultan sin duda de enorme relevancia para la más profunda comprensión de la historia léxica del español (tabla 3).

Americanismo puro (AP) → Americanismo de frecuencia (AF) → Voz General (VG)		
<i>atol, calpul, chachaguatate, chachigüite, chapulín, chichigua, jioté, jiquilite, ocote, pacaya, tecolote</i>	<i>aguacate, cacao, chácara, nagua, papaya, yuca</i>	<i>jícara, nopal</i>

Tabla 3. Estatus de (algunos) indigenismos del corpus (desamericanización)

5. Unas primeras conclusiones

De este modo, esta rápida revisión de los indigenismos que aparecen en los *Vicios del lenguaje y provincialismos de Guatemala* ofrece ya una serie de conclusiones que responden a los objetivos planteados al inicio de estas páginas y permiten, en consecuencia, no solo comprender mejor el valor que Batres Jáuregui y la lexicografía centroamericana decimonónica concede a los préstamos de origen amerindio, sino también aportar datos de cierta relevancia para poder conocer más profundamente, y desde el punto de vista lexicológico, la —por el momento poco atendida— historia léxica de la variedad guatemalteca del español.

Por lo que se refiere a la primera cuestión, es importante mencionar que el autor de la obra, a pesar de constituir un claro exponente del movimiento purista de su época que busca censurar “todo tipo de expresión dialectal que atentara contra la unidad lingüística” (Quesada Pacheco, 2008, p. 151), no muestra

en principio una actitud tan negativa para con el indigenismo como el título de su trabajo podría dar a entender; al contrario, sus páginas ofrecen en realidad una clara defensa de estas voces con razones de muy diversa índole —tales como su uso más o menos generalizado en el continente, su idoneidad para expresar nuevos conceptos o su importancia para la identidad lingüística americana—, así como un alegato a favor de su incorporación al cuerpo del idioma por lo que supone de enriquecimiento de una lengua que necesariamente debe ir más allá de los límites de la norma peninsular. No sorprende, por tanto, que las entradas que hacen referencia a estos elementos se desarrollen, más que como una crítica, como una explicación en la que se aportan datos muy variados —sobre todo, la etimología, la extensión geográfica y el significado preciso del vocablo— y en la que a menudo se corrigen las informaciones facilitadas por el *Diccionario* de la Real Academia Española, en lo que supone una reivindicación explícita y una observación de las hablas americanas que parecen preludiar el paso de lo preceptivo a lo descriptivo que se produce a inicios del siglo xx en la lexicografía de América Central (Quesada Pacheco, 2003, p. 550).

Por otro lado, resulta lógico pensar que esta visión positiva del indigenismo que se acaba de mencionar tendrá como consecuencia la aparición de numerosos vocablos de este origen en sus páginas, y es precisamente esto lo que se constata a la luz de los datos cuantitativos recogidos en este estudio: en efecto, la lectura cuidadosa del leuario de Batres Jáuregui arroja un total de 230 entradas indudables —197 voces simples y 33 derivadas— de esta naturaleza, a las que probablemente se puedan añadir algunas más cuya etimología no aparece en las fuentes lexicográficas consultadas. Si se tiene en cuenta que la obra analizada se compone de 1800 entradas (Quesada Pacheco, 2003, p. 550), se puede concluir que el indigenismo representa al menos el 12.77% de todas ellas, un porcentaje notablemente elevado que, como se señaló ya anteriormente, constata la trascendencia de este texto para el estudio del préstamo amerindio empleado en la Guatemala decimonónica y, por tanto, para la reconstrucción diacrónica del vocabulario diatópicamente restringido que dota de personalidad al habla de este país.

Ahora bien, la trascendencia que se acaba de mencionar más arriba no se queda en lo puramente cuantitativo, sino que es evidente también desde otros puntos de vista. Así, el análisis de los orígenes etimológicos de los indigenismos no solo refuerza algunas ideas ya conocidas sobre la importancia de determinadas lenguas amerindias en el enriquecimiento léxico del español guatemalteco —la importancia del náhuatl, el escaso peso del maya o la presencia de vocablos antillanos—, sino que además constata cuestiones de interés cuyo análisis habrá que realizar más adelante, y entre las cuales destacan sin duda dos: por un lado, la continuidad fundamental que en las diversas tendencias muestra este corpus decimonónico respecto a otros momentos históricos; por otro, la abundancia y dispersión geográfica de los sistemas lingüísticos que aportan voces al español guatemalteco, lo que revela la complejidad que esconde la configuración léxica de las hablas americanas, resultado no solo de la convivencia con los idiomas autóctonos, sino también de los trasvases léxicos e influencias mutuas que a lo largo de la historia se establecen entre las diversas regiones del continente. Por otro lado, la revisión de los campos léxicos a los que se incorporan estos elementos resulta también de interés por dos motivos: por un lado, porque su extensión por ámbitos de la realidad muy variados demuestra que la entrada de indigenismos supone realmente una completa reorganización del vocabulario del español americano; por otro, porque su concentración mayoritaria en ciertas esferas como la agricultura, los enseres o la fauna y la flora —que responde a la apropiación de realidades americanas por parte de la sociedad hispánica que se configura en el Nuevo Mundo— sigue las tendencias ya apuntadas en otros estudios y demuestra, una vez más, la estrecha relación existente entre los procesos históricos y los cambios lingüísticos.

Pasando ahora a la perspectiva de uso y, con ello, al peso de estos elementos en la conformación de la identidad léxica de la variedad guatemalteca del español, es importante mencionar las diferencias que, respecto a la definición aquí utilizada del *americanismo* (Ramírez Luengo, 2017), presentan estos elementos, los cuales no solo se distribuyen entre los tres subtipos existentes dentro de este concepto —a saber, puros (*calpul*, *chapulín*,

pacaya), semánticos (*guanaco*) y de frecuencia (*cacao*, *chácara*, *yuca*)—, sino que en ocasiones se han transformado incluso en voces generales (*jícara*, *nopal*), lo que reafirma el carácter dinámico que identifica al americanismo y constata una vez más la existencia de procesos de desamericanización de los indigenismos (Ramírez Luengo, 2017b), aún poco analizados diacrónicamente. Por último, y desde lo estrictamente cronológico, la obra de Batres Jáuregui resulta también de notable relevancia, pues —a pesar de su modernidad— sus páginas aportan no solo la primera datación absoluta y/o centroamericana de numerosos vocablos (por ejemplo, *bucul*, *canche*, *güicoy*, *pupusa*, o *tetunte* en el caso de los primeros; *canchalagua*, *chipilín*, *huisache*, *nigua* o *tepocate* en los segundos), sino también la primera aparición histórica de voces solo recogidas en sincronía (*cantil*, *chibola*, *jute*, *pilixte*, *tihuilote*, *zute*) o incluso ejemplos de unidades léxicas con muy escasa atestiguación histórica (*calzonte*, *chajal*, *chocoyo*, *copalchi*), todo lo cual transforma a los *Vicios del lenguaje* en una obra fundamental para el más correcto rastreo diacrónico del aporte de las lenguas autóctonas americanas al español de Guatemala.

En definitiva, no cabe duda de que este primer acercamiento a la obra de Batres Jáuregui revela la enorme importancia que, desde diferentes puntos de vista —tanto lo historiográfico como lo más estrictamente lexicológico—, posee para la historia del español centroamericano el trabajo de este autor. Es necesario, por tanto, seguir indagando desde un punto de vista filológico las páginas de sus *Vicios del lenguaje* y *provincialismos de Guatemala* y desarrollar estudios que, partiendo de ellas, analicen cuestiones tan variadas como sus ideas lingüísticas y las fuentes en las que basa sus apreciaciones, los rasgos fónicos y morfosintácticos que describe —auténtico diagnóstico sociodialectal de su época— o, ya en el vocabulario, los arcaísmos, galicismos y anglicismos que documenta, pues salta a la vista que es aún mucha, y de gran relevancia, la información que esconde este texto para poder avanzar de forma segura en el estudio del devenir diacrónico que, con el paso del tiempo, termina por configurar la variedad de español que identifica hoy a los guatemaltecos.

Referencias

- BATRES JÁUREGUI, A. (1892). *Vicios del lenguaje y provincialismos de Guatemala*, Ciudad de Guatemala: Tipografía Nacional.
- BUESA OLIVER, T., y ENGUITA UTRILLA, J. M. (1992). *Léxico del español de América: su elemento patrimonial e indígena*. Madrid: MAPFRE.
- BUZEK, I., y ŠINCOVÁ, M. (2014). Introducción: *Una cercana sincronía opaca: estudios sobre el español del siglo XIX* (parte primera). *Études Romanes de Brno*, 36(1), 7-10.
- BVFE. Alvar Ezquerro, M. (2022). *Biblioteca Virtual de la Filología Española (BVFE). Directorio bibliográfico de gramáticas, diccionarios, obras de ortografía, ortología, prosodia, métrica, diálogos e historia de la lengua*. Recuperado de <<http://www.bvfe.es>>.
- CASAUS ARZÚ, M. E. (1999). Los proyectos de integración social del indio y el imaginario nacional de las elites intelectuales guatemaltecas, siglos XIX y XX. *Revista de Indias*, 59(217), 775-813.
- COMPANY, C. (2010). Introducción. En Academia Mexicana de la Lengua, *Diccionario de Mexicanismos* (pp. XV-XXIII). México: Siglo XXI.
- COMPANY, C. (2017). El siglo XIX en la periodización sintáctica de la lengua española. En E. Carpi y R. M. García Jiménez (Eds.), *Herencia e innovación en el español del siglo XIX* (pp. 75-101). Italia: Pisa University Press.
- CORDE. Real Academia Española (2022). *Banco de datos CORDE. Corpus diacrónico del español*. Recuperado de <<http://www.rae.es>>.
- CORDIAM. Academia Mexicana de la Lengua (2022). *Corpus Diacrónico y Diatópico del Español de América*. Recuperado de <<http://www.cordiam.org>>.
- DAMER. Asociación de Academias de la Lengua Española (2010). *Diccionario de Americanismos*. Madrid: Santillana. Recuperado de <<https://www.asale.org/damer/>>.
- DCECH. Corominas, J., y Pascual, J. A. (1980-1991). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.

- DHECR. Quesada Pacheco, M. A. (1995). *Diccionario histórico del español de Costa Rica*. Costa Rica: Editorial Universidad Estatal a Distancia.
- DLE. Real Academia Española (2014). *Diccionario de la Lengua Española*. Madrid: Espasa-Calpe. Recuperado de <<https://dle.rae.es>>.
- DYM, J., y HERRERA MENA, S. A. (2014). *Centroamérica durante las revoluciones atlánticas: el vocabulario político, 1750-1850*. El Salvador: IEESFORD Editores.
- ERICKSON, M. E. (1942). Antonio Batres Jauregui: Guatemalan Critic. *Hispania*, 25(3), 343-350.
- GARCÍA GIRÁLDEZ, T. (2009). El pensamiento político liberal centroamericano del siglo XIX: José Cecilio del Valle y Antonio Batres Jáuregui. *Revista Complutense de Historia de América*, 35, 23-45.
- HERRANZ, A. (2017). *Procesos de nahuatlización y nahuatlismos de uso en Honduras*. Tegucigalpa: Universidad Nacional Autónoma de Honduras.
- HERRERA PEÑA, G. (2021). Rasgos del español guatemalteco en dos obras atribuidas a Sor Juana de Maldonado y Paz. *Diseminaciones*, 4(7), 15-32.
- HERRERA PEÑA, G. (2021b). El español literario en Guatemala durante las primeras décadas del siglo XIX. En G. Cruz Volio, A. Sánchez Mora y J. L. Ramírez Luengo (Eds.), *La palabra olvidada. Lengua y literatura en Centroamérica (siglos XVI-XIX)* (pp. 131-170). Costa Rica: Editorial Encino.
- LEXHISP. Boyd-Bowman, P. (2003). *Léxico hispanoamericano. 1493-1993*. Nueva York: Hispanic Seminary of Medieval Studies. Recuperado de <<http://textred.spanport.lss.wisc.edu>>.
- MARTÍNEZ PUGA, C. de la L. (2021). El léxico de los tejidos en la Guatemala del siglo XVIII: notas sobre el inventario de la visita del obispo Fray Mauro de Larreátegui (1707). En J. L. Ramírez Luengo (Ed.), *Estudios sobre la historia léxica del español de América* (pp. 117-124). España: Universidad de Jaén.
- MEJÍAS, H. A. (1980). *Préstamos de las lenguas indígenas en el español americano del siglo XVII*. México: Universidad Nacional Autónoma de México.

- MONTERO LAZCANO, M. Y. (2020). Indigenismos en el discurso médico de Guatemala del siglo XVIII: el caso de la *Instrucción sobre el modo de practicar la inoculación de las viruelas* de José Felipe Flores. *Études Romanes de Brno*, 41(2), 41-51.
- MORÍNIGO, M. A. (1998). *Nuevo Diccionario de Americanismos e Indigenismos*. Buenos Aires: Claridad.
- POLO CANO, N. (2005). Algunos indigenismos léxicos en el español de Guatemala del siglo XVIII. *Res Diachronicae Virtual 4: El contacto de lenguas*, 185-202. Recuperado de <<http://www.resdi.net>>.
- QUESADA PACHECO, M. A. (2003). Lexicografía del español de América Central: esbozo histórico. En *Lexicografía y lexicología en Europa y América: Homenaje a Günter Haensch* (pp. 547-558). Madrid: Gredos.
- QUESADA PACHECO, M. A. (2008). El español de América Central ayer, hoy y mañana. *Boletín de Filología de la Universidad de Chile*, 43, 145-174.
- RAMÍREZ LUENGO, J. L. (2007). *Breve historia del español de América*. Madrid: Arco Libros.
- RAMÍREZ LUENGO, J. L. (2011). *La lengua que hablaban los próceres. El español de América en la época de las Independencias*. Buenos Aires: Voces del Sur.
- RAMÍREZ LUENGO, J. L. (2012). El español del siglo XIX. O la historia de un abandono. En J. L. Ramírez Luengo (Ed.), *Por sendas ignoradas. Estudios sobre el español del siglo XIX* (pp. 7-10). España: Axac.
- RAMÍREZ LUENGO, J. L. (2012b). El léxico en los procesos de dialectalización del español americano: el caso de la Bolivia andina. *Cuadernos del Instituto de Historia de la Lengua Española*, 7, 393-404.
- RAMÍREZ LUENGO, J. L. (2015). El americanismo y la historia: algunas pautas diacrónicas de creación de americanismos (léxicos). *Miradas*, 13, 114-123.
- RAMÍREZ LUENGO, J. L. (2017). Aspectos metodológicos para el estudio histórico del léxico americano: conceptos, ejemplificación y tareas para el futuro. *Moenia*, 23, 603-619.

- RAMÍREZ LUENGO, J. L. (2017b). Los corpus lingüísticos en la historia del léxico: algunos datos sobre la generalización de los indigenismos antillanos en el español de España. *Études Romanes de Brno*, 38(2), 101-111.
- RAMÍREZ LUENGO, J. L. (2019). Algunas notas sobre el léxico del mestizaje en la Guatemala del siglo XVIII (1690-1810). *RI-LEX. Revista de investigaciones léxicas*, 2(1), 5-24.
- RAMÍREZ LUENGO, J. L. (2019b). Los indigenismos léxicos en la *Descripción geográfico-moral* del arzobispo Cortés y Larraz (1770). Los datos salvadoreños. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 45(2), 249-265.
- RAMÍREZ LUENGO, J. L. (2020). El léxico de la medicina en la Guatemala del siglo XVIII: algunas notas sobre la *Instrucción sobre el modo de practicar la inoculación de las viruelas* de José Felipe Flores (Ciudad de Guatemala, 1794). *Études Romanes de Brno*, 41(2), 27-39.
- RAMÍREZ LUENGO, J. L. (2021). La americanización léxica del español guatemalteco de la Ilustración a partir de los *Apuntamientos sobre la agricultura y el comercio del Reyno de Guatemala* (1811). En *Publicación conmemorativa Bicentenario de la Independencia 1821-2021. La lengua española en Guatemala* (pp. 67-111). Guatemala: Academia Guatemalteca de la Lengua.
- RAMÍREZ LUENGO, J. L. (2021b). Los indigenismos en el discurso médico de la Guatemala decimonónica: una aproximación a partir del *Método preservativo y curativo de la cólera* (1833). *Ciencias Sociales y Humanidades. Revista Centroamericana de Investigación y Postgrado*, 8(2). Recuperado de <<http://revistas.usac.edu.gt/index.php/csh/article/view/1068>>.
- RAMÍREZ LUENGO, J. L. (en prensa). *Pintas negras o moradas que llaman vulgarmente alfombrilla*: las voces coloquiales en el discurso médico guatemalteco del siglo XVIII. En M. F. Gonçalves (Ed.), *A linguagem médica no século XVIII: corpora e humanidades digitais*. Portugal: Universidad de Évora.
- RAMÍREZ LUENGO, J. L. (en prensa b). La metáfora en la conformación del léxico médico de la Guatemala ilustrada: algu-

nos ejemplos de la *Instrucción sobre el modo de practicar la inoculación de las viruelas* de José Felipe Flores (Ciudad de Guatemala, 1794). *Cuadernos AISPI. Estudios de lenguas y literaturas hispánicas*.

RAMÍREZ LUENGO, J. L. (en prensa c). Los americanismos léxicos en el español de la Guatemala ilustrada: una aproximación desde los *Apuntamientos sobre la agricultura y el comercio del Reyno de Guatemala* (1811). En R. Grajales y L. A. Ciro (Eds.), *Estudios lingüísticos e interdisciplinarios en Latinoamérica*. Colombia: Universidad del Valle.

RAMÍREZ LUENGO, J. L. (en prensa d). Las *Observaciones* de Henríquez Ureña (1921), más allá de la dialectología: algunas notas desde/para la historia de la lengua. *Boletín de Filología de la Universidad de Chile*.

RAMÍREZ LUENGO, J. L., AQUINO MELCHOR, S. A., y RAMÍREZ VÁZQUEZ, J. J. (2017). Indigenismos léxicos en el español guatemalteco del siglo XVIII. *Lengua y habla*, 21, 111-125.

RAMÍREZ LUENGO, J. L., SAN MARTÍN GÓMEZ, J. A. (2020). Procesos de americanización léxica en la historia del español salvadoreño: la carta de Diego García de Palacio como ejemplo (siglo XVI). *Revista de Filología de la Universidad de La Laguna*, 41, 183-195.

RIVAS HIDALGO, J. D. (2022). *Las ideologías dominantes en la lexicografía centroamericana (1892-1910)*. El Salvador y Guatemala. Tesis de maestría, Universidad de El Salvador, San Salvador.

SANTAMARÍA, F. J. (1959). *Diccionario de mejicanismos*. México: Porrúa.